A Probability-guided Sampler for Neural Implicit Surface Rendering

Gonçalo Dias Pais [1,2], Valter Piedade [2], Moitreya Chatterjee [1], Marcus Greiff [3], and Pedro Miraldo [1]

Mitsubishi Electric Research Labs (MERL)[1], Instituto Superior Técnico[2], Toyota Research Institute[3]

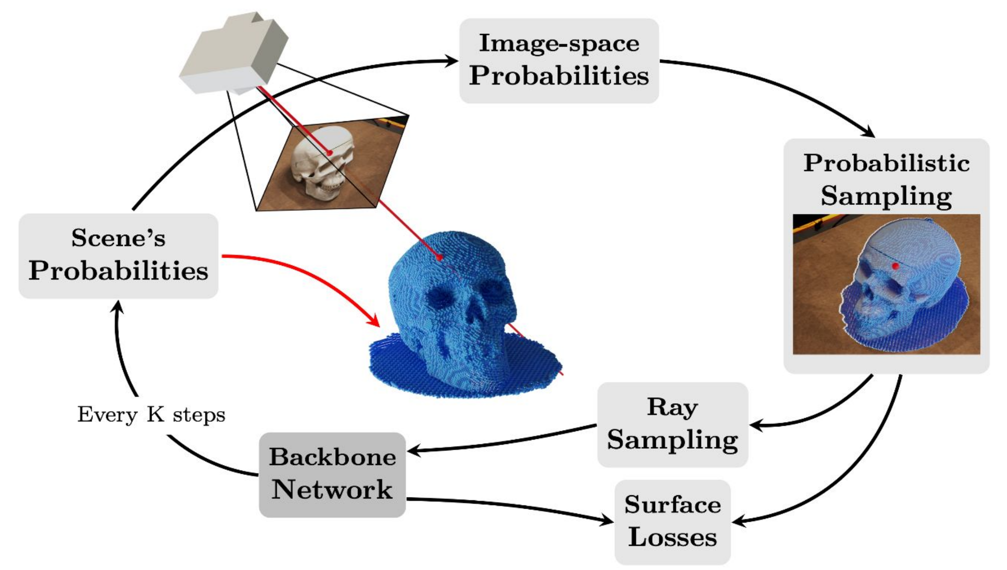
EUROPEAN CONFERENCE ON COMPUTER VISION

MITSUBISHI ELECTRIC RESEARCH LABORATORIES, INC

Motivation and Contributions

- Neural implicit surfaces methods extend NeRF and allow to obtain finer details and higher resolution 3D surfaces while training for image synthesis;
- Previous works show that depth regularization improves neural rendering and 3D reconstruction, but these rely on additional data, like SfM points.

Can the implicit surface representation guide the sampling of rays and points during training for accurate 3D reconstruction and rendering?



Neuralangelo

Neuralangelo w/ Our Sampler

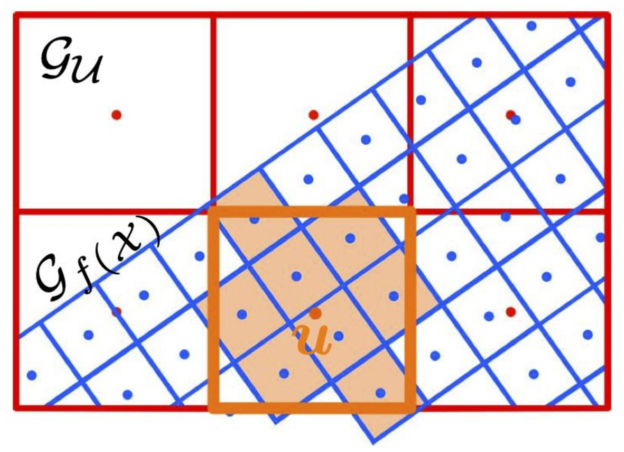
Pipeline

1) Obtain the scene's SDF $S(\mathbf{x})$ from the implicit surface as probabilities $\phi_s(S(\mathbf{x}))$;
2) For each camera, we transform the scene's probabilities to the proposed image space at every K iteration;
3) At each iteration, sample the image space:
  - Each sample includes: image pixel and estimated depth;
  - The sampled depth is used to regularize the model;
  - The method is agnostic to the backbone. We used **Neuralangelo**[1] and **NeuS**[2] in our experiments.

Probability-guided Sampler

Interpolation

- We approximate $p(\mathbf{u})$ as the Riemann integral of all cells of $\mathcal{G}_{f(\mathcal{X})}$ in $\mathbf{u}$:
  - $\mathcal{G}_{\mathcal{U}}$: 3D image space in red;
  - $\mathcal{G}_{f(\mathcal{X})}$: Scene grid in the image space in blue;
  - $\mathbf{u}$: cell $\square \in \mathcal{G}_{\mathcal{U}}$ in orange;
  - $\mathbf{v} = f(\mathbf{x})$, η is the depth of $\mathbf{v}$.



$$p(\mathbf{u}) = \sum_{\bullet \in \square} \eta_{\bullet}^{-2} p(\boldsymbol{x}_{\bullet})$$

View Dependency

- Since $p(\mathbf{u})$ does not account for occlusions, sampling a projection ray based on the object's geometry alone can result in many occluded samples;
- The view-dependent probability $\widetilde{p}(\mathbf{u}_i)$, where for $\mathbf{u}_i = [u, v, \lambda_i]^T$ considers the transmittance T_i along the depth λ_i of $[u, v]$ and $\sigma_i = p(\mathbf{u}_i)$:

$$\widetilde{p}(\mathbf{u}_i) = \sigma_i T_i = p(\mathbf{u}_i) e^{-\sum_{k=0}^{i} p([u,v,\lambda_k]^T)}$$

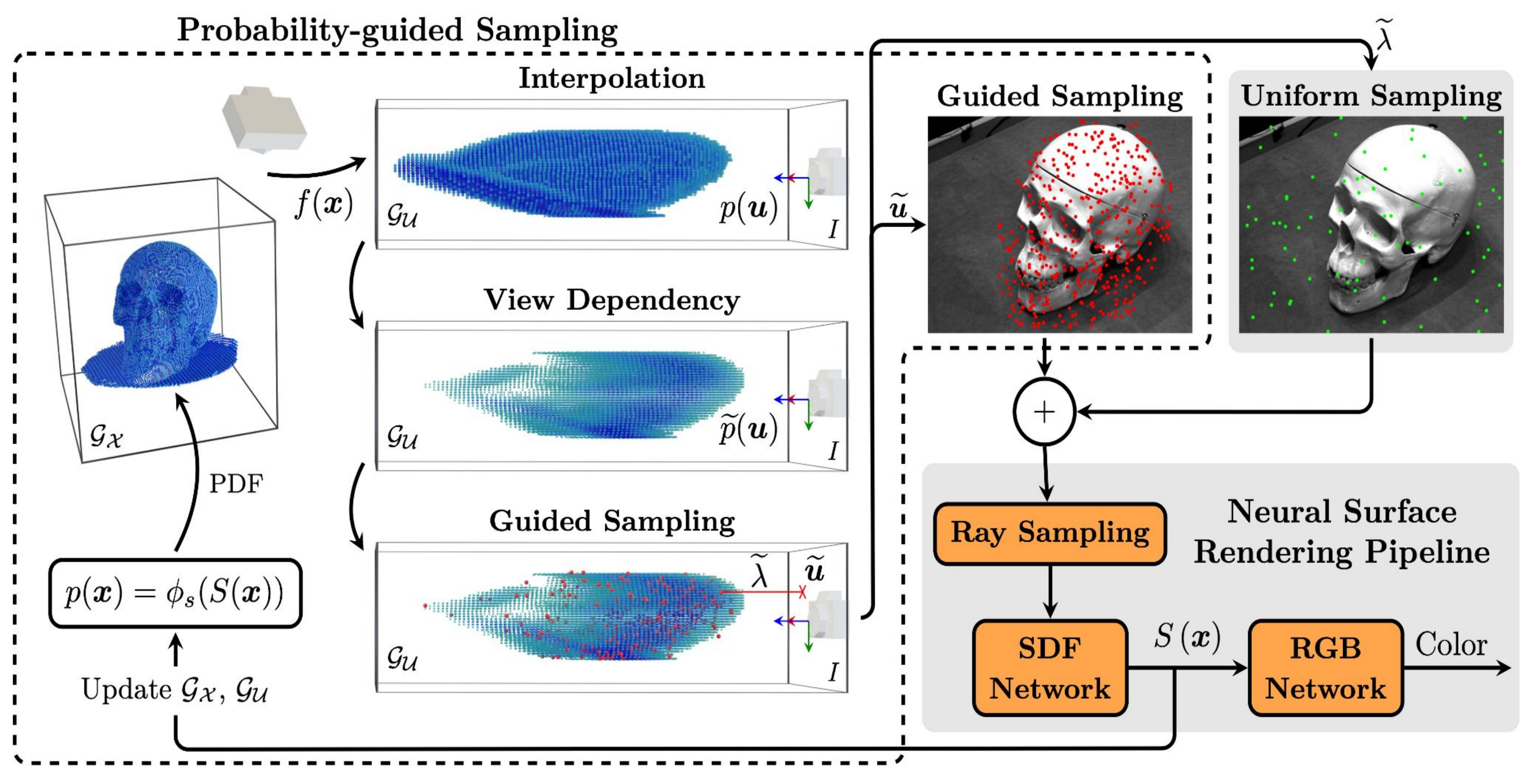
Probability-guided Sampling

- We combine two image sampling strategies:
  - *(i)* Sampling using the view-dependent PDF $\widetilde{p}(\mathbf{u})$:
    - Conditional importance sampling in the 3-axis;
    - Targets the surfaces in the scene.
  - *(ii)* Sampling uniformly on the image:
    - Allows background rendering;
- Initial training stages focus on *(i)*. Then, we increase samples of *(ii)*, while keeping the same number sampled rays.

Surface Reconstruction Losses

- In addition to the losses of each backbone, we include a surface loss for:
  - *(i)* Points near the surface: L^{Near};
  - *(ii)* Points within the empty ray space: L^{Empty};
  - *(iii)* Points belonging to background rays: L^{Bg};
- The total surface loss is computed as:

$$L^{\text{Surf}} = \lambda_1 L^{\text{Near}} + \lambda_2 (L^{\text{Bg}} + L^{\text{Empty}})$$

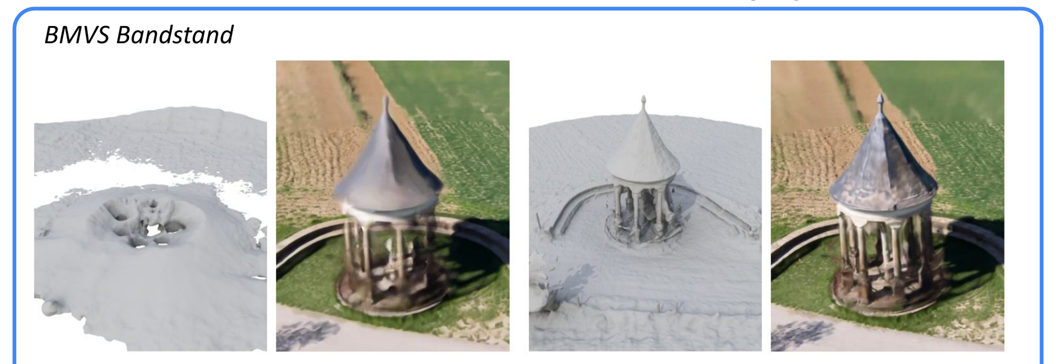


Results

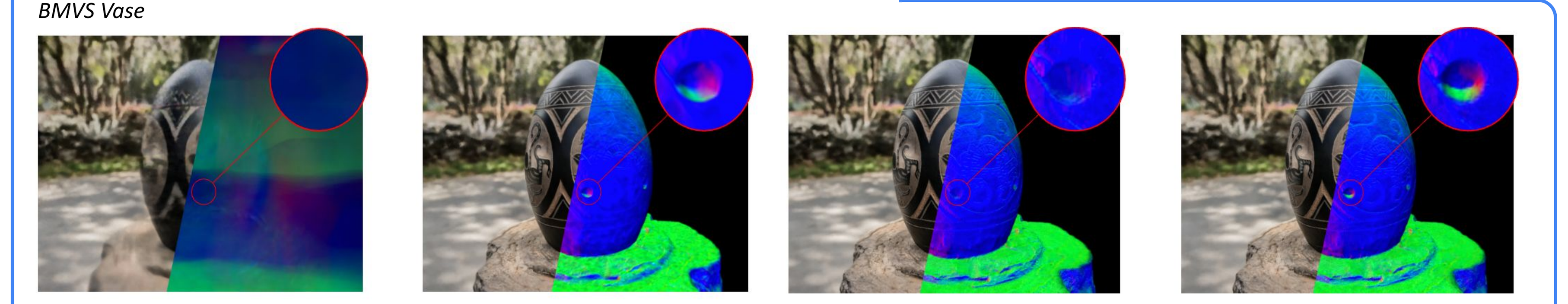
Better accuracy

(a) Neuralangelo (b) Neuralangelo + Ours

Challenging reconstructions

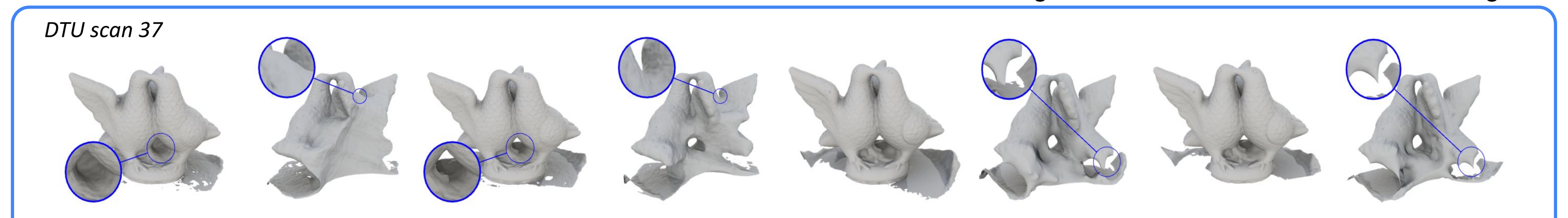


(a) NeuS (b) NeuS + Ours



(a) NeuS (b) NeuS + Ours (c) Neuralangelo (d) Neuralangelo + Ours

Higher reconstruction in less viewed regions

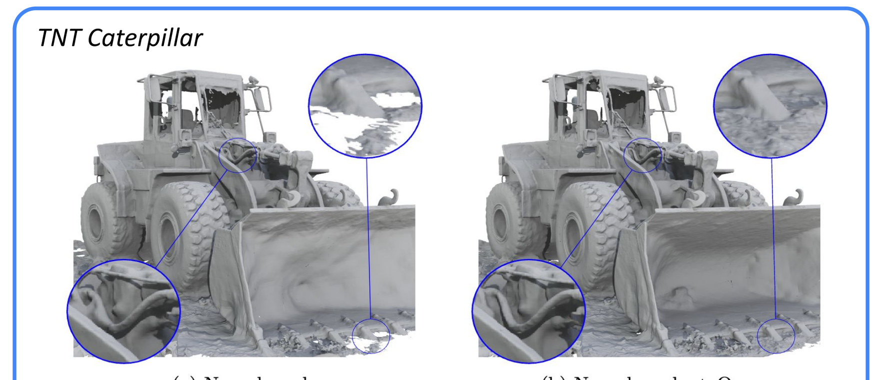


(a) NeuS (b) NeuS + Ours (c) Neuralangelo (d) Neuralangelo + Ours

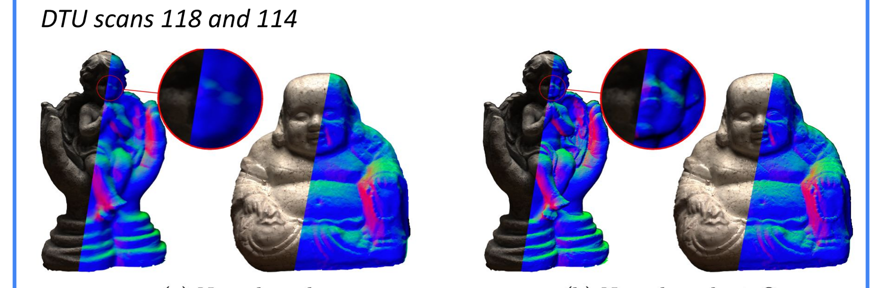
DTU Quantitative Results

			24	37	40	55	63	65	69	83	97	105	106	110	114	118	122	**Mean**
PSNR ↑	Unmasked	NeRF [124]	26.24	25.74	26.79	27.57	31.96	31.50	29.58	32.78	28.35	32.08	33.49	31.54	31.00	35.59	35.51	30.65
		VolSDF[211]	26.28	25.61	26.55	26.76	31.57	31.50	29.38	33.23	28.03	32.13	33.16	31.49	30.33	34.90	34.75	30.38
		RegSDF† [224]	24.78	23.06	23.47	22.21	28.57	25.53	21.81	28.89	26.81	27.91	24.71	25.13	26.84	21.67	28.25	25.31
		NeuS [191]	23.85	27.63	27.16	29.4	32.71	33.1	30.58	34.25	29.97	33.69	35.34	32.81	31.96	36.72	37	31.74
		NeuS + Ours	28.28	28.1	28.16	24.71	33.1	33.97	29.59	33.25	30.35	33.61	35.66	32.97	32.29	37.15	35.56	31.78
		Neuralangelo [103]	30.64	27.78	32.70	**34.18**	35.15	**35.89**	31.47	**36.82**	30.13	**35.92**	36.61	32.60	31.20	**38.41**	**38.05**	33.84
		Neuralangelo + Ours	**33.73**	**30.36**	**33.55**	34.06	**35.22**	34.64	**32.49**	33.2	**31.93**	34.17	**37.64**	**35.3**	**34.01**	38.04	37.87	**34.41**
	Masked	NeuS [191]	28.93	28.29	27.53	30.57	36.48	36.48	31.83	40.59	31.26	37.19	36.87	33.9	32.65	39.63	40.88	34.21
		NeuS + Ours	28.98	29.22	28.66	25.22	37.24	38.73	30.77	42.47	32.34	37.5	37.49	34.34	33.11	40.84	38.45	34.36
		Neuralangelo [103]	**35.21**	31.76	35.12	38.16	41.17	40.46	34.39	44.22	34.09	40.8	40.8	37.24	34.92	42.36	43.56	38.28
		Neuralangelo + Ours	35.13	**32.86**	**35.2**	**38.51**	**41.41**	**41**	**34.51**	**44.93**	**35.64**	**41.13**	**40.95**	**37.78**	**35.26**	**43.3**	**44.59**	**38.81**
Chamfer (mm) ↓		NeRF [124]	1.90	1.60	1.85	0.58	2.28	1.27	1.47	1.67	2.05	1.07	0.88	2.53	1.06	1.15	0.96	1.49
		VolSDF[211]	1.14	1.26	0.81	0.49	1.25	0.70	0.72	1.29	1.18	0.70	0.66	1.08	0.42	0.61	0.55	0.86
		HF-NeuS [197]	0.76	1.32	0.70	0.39	1.06	0.63	0.63	1.15	1.12	0.80	0.52	1.22	0.33	0.49	0.50	0.77
		RegSDF† [224]	0.60	1.41	0.64	0.43	1.34	0.62	0.60	0.90	0.92	1.02	0.60	0.59	0.30	0.41	0.39	0.72
		NeuralWarp [42]	0.49	0.71	0.38	0.38	**0.79**	0.81	0.82	**1.20**	1.06	**0.68**	0.66	**0.74**	0.41	0.63	0.51	0.68
		NeuS [191]	0.77	0.78	5.82	0.50	1.39	1.76	1.06	4.01	1.47	0.77	0.64	1.29	0.34	0.56	0.53	1.30
		NeuS + Ours	1.08	0.74	1.27	2.43	1.05	1.05	1.66	1.32	2.1	0.79	0.6	1.07	0.32	0.4	2.08	1.2
		Neuralangelo [103]	**0.37**	0.72	0.35	0.35	0.87	**0.54**	**0.53**	1.29	0.97	0.73	**0.47**	**0.74**	0.32	0.41	0.43	0.61
		Neuralangelo + Ours	0.39	**0.68**	**0.32**	**0.33**	0.87	0.58	**0.53**	1.3	**0.93**	0.70	0.5	**0.74**	**0.31**	**0.37**	**0.38**	**0.6**

Sharper details



(a) Neuralangelo (b) Neuralangelo + Ours



(a) Neuralangelo (b) Neuralangelo + Ours

[1] Li *et al.* "Neuralangelo: High-fidelity neural surface reconstruction". CVPR. 2023.
[2] Wang *et al.* "NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction". NIPS. 2021.

{goncalo.pais,valter.piedade}@tecnico.ulisboa.pt,chatterjee@merl.com,marcus.greiff@tri.global,miraldo@merl.com

Pais and Piedade were partially supported by LARSyS, Portuguese "Fundação para a Ciência e a Tecnologia" (FCT) funding (DOI: 10.54499/LA/P/0083/2020, 10.54499/UIDP/50009/2020, and 10.54499/UIDB/50009/2020). Pais was also partially supported by FCT grant PD/BD/150630/2020. Pedro Miraldo and Moitreya Chatterjee are exclusively supported by MERL.